Impact of CMOS Technologies scaling in Leakage Reduction Techniques

Saad Arrabi, Taeyoung Kim

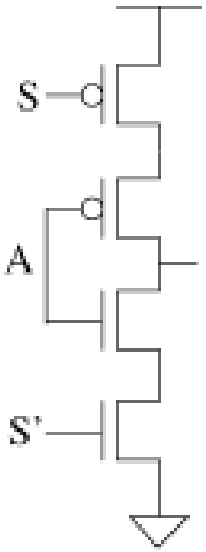
OUTLINE

- INTRODUCTION
- BACKGROUND
- METHODOLOGIES
- SIMULATION RESULTS
- CONTRIBUTION
- FUTURE WORK

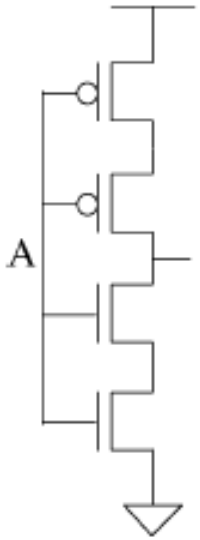
INTRODUCTION

- MOTIVATION
  - Power consumption is the major concern in CMOS semiconductor industries. Leakage is one of the big part of power consumption.
  - Off current leakage is being more serious problem as the technology size shrinks.
- GOALS
  - Explore what conventional techniques for leakage reduction is used.
  - Investigate how effectively they can cope with new technology scaling.

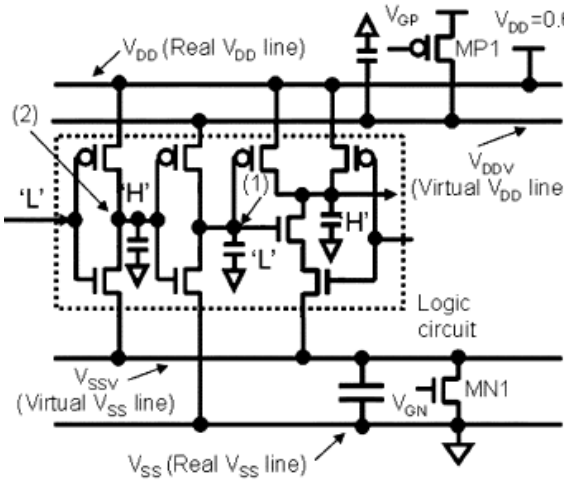
BACKGROUND



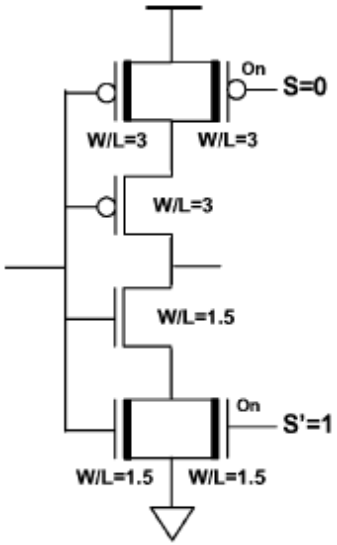
Sleep



Stack



Zigzag CMOS



Sleepy Stack

Sleep

- NMOS Sleep transistor usually used
  - To reduce resistance of “on” status
- Pros
  - Reduction dynamic and leakage power
- Cons
  - Reduce performance and requires additional area
  - Energy is consumed by charging and discharging the virtual rails
- Knob
  - Footer vs. Header
  - Width
  - Sharing vs. Non-sharing Sleep TR
    - Sharing will decrease the delay

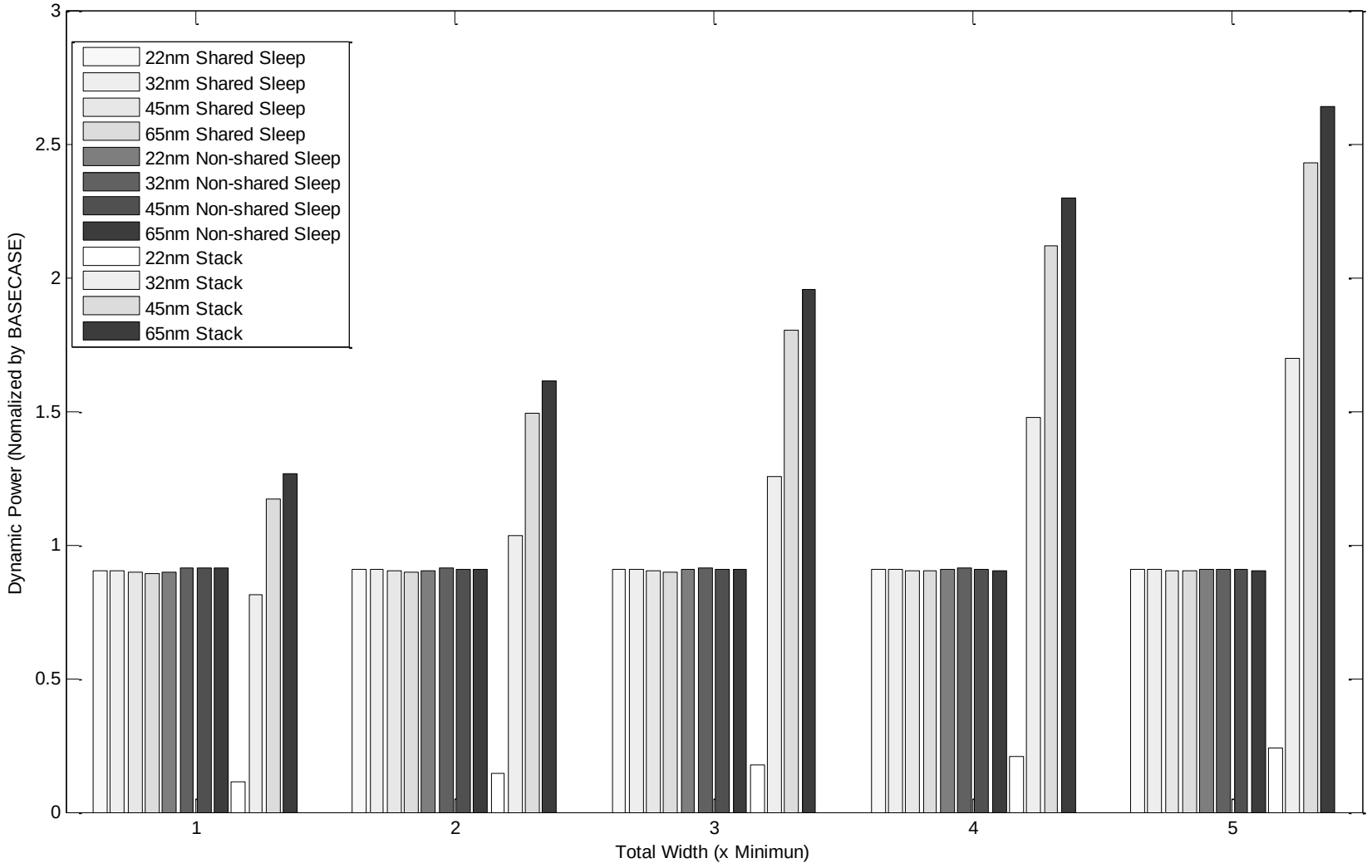
STACK

- Doesn’t need to go to sleep mode
- Introduces new capacitance and delay
- Knobs
  - Width: more width, more leakage, more area, less delay
  - Placement: Headers only, some inputs.
    - Different leakage reduction and less area

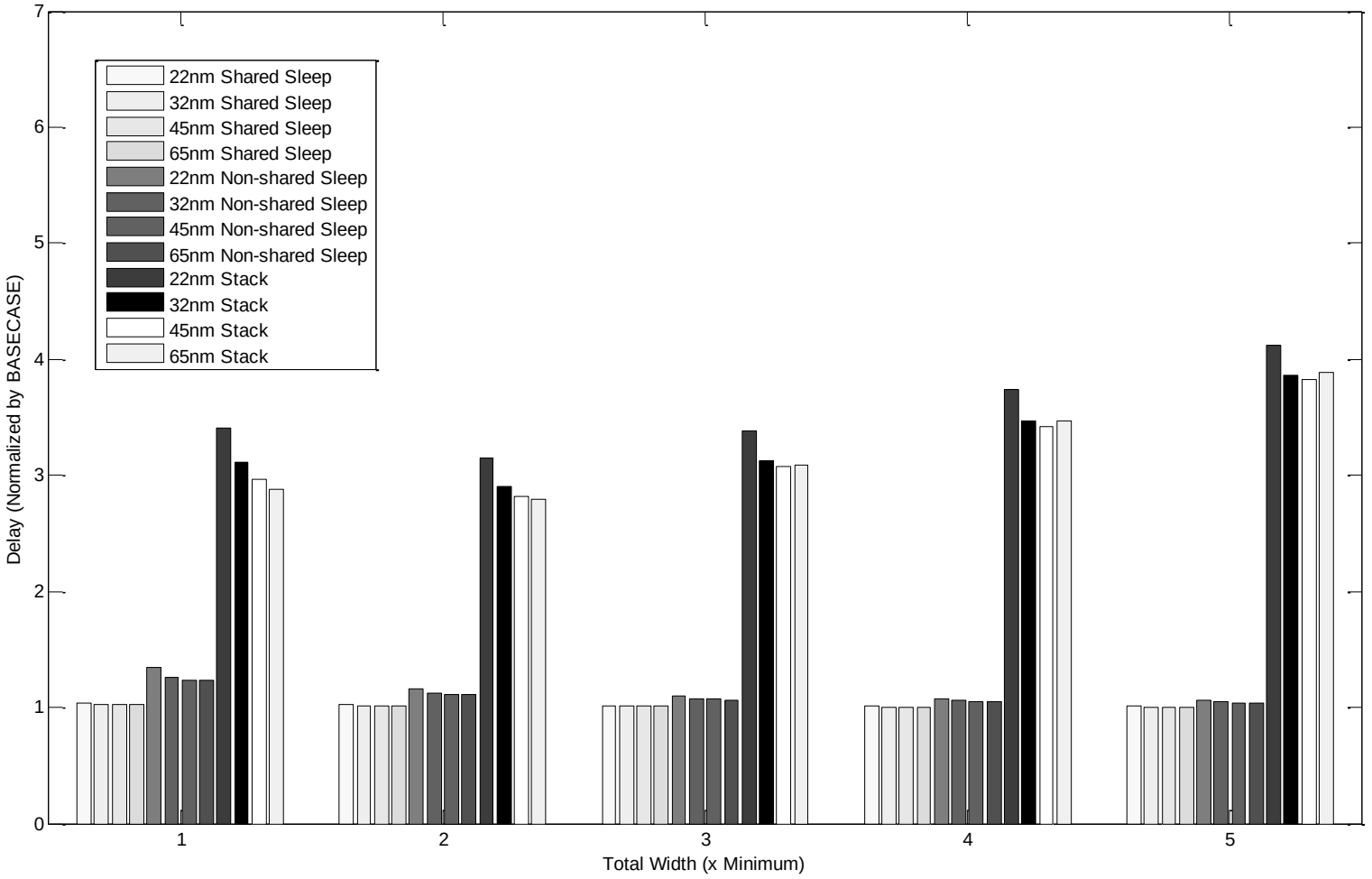
Methodologies

- RANDOM CIRCUIT
  - Simulate glitching
  - Large enough to be between 2 registers stages
  - Not fully accurate, but good enough
- DELAY
  - Dynamic Delay, 50% of input to 50% of output
  - Wake up Delay, 50% of input signal to 90% of rail voltage
- ENERGY
  - Dynamic Energy, averaged between 0->1, 1->0
  - Leakage Energy, averaged between static 0, 1
  - Wake up Energy, same as delay

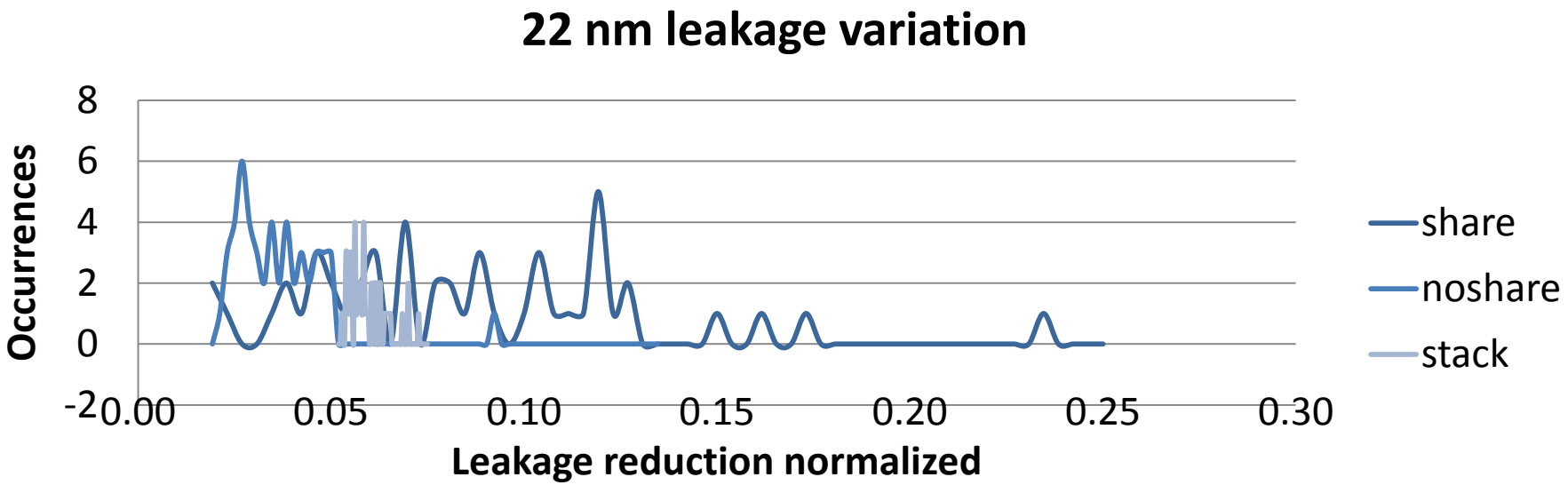
Dynamic Power vs Total Widths

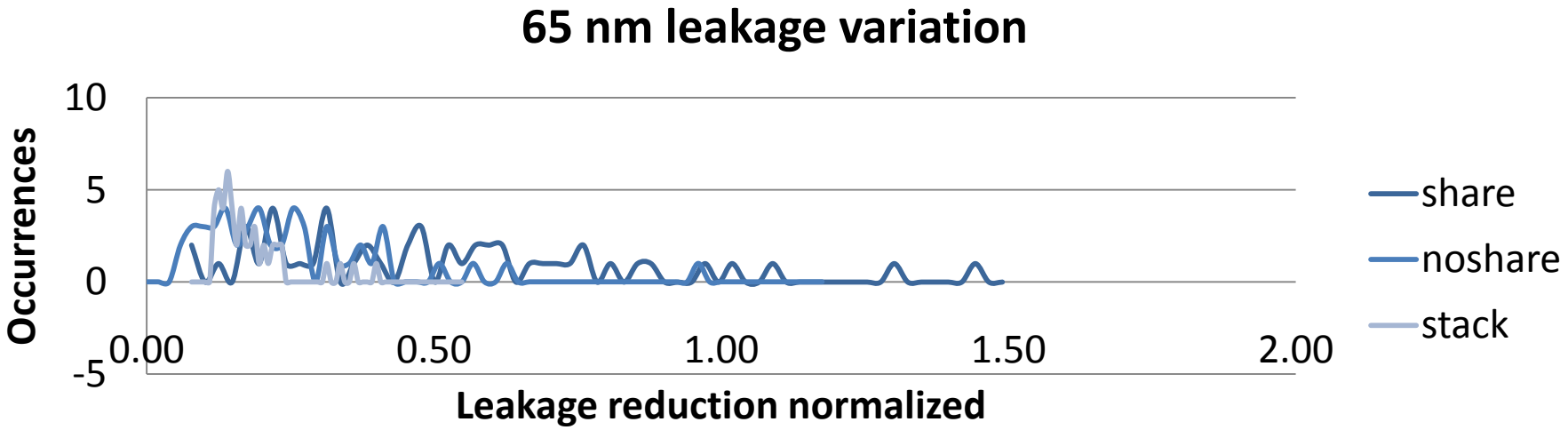


Dynamic Delay vs. Total Widths



Leakage reduction variation





Comparison

Metric	Stack	Sleep
Added delay	High	Medium
Added energy	High	Low
Leakage in active mode	Medium	None
Leakage in sleep	Medium	High
Complexity	Simple	Complex
Area	Big area	Medium area

Contribution

- Showed the effectiveness of some techniques for future technologies
- Provide a base work for knob tweaking for leakage reduction techniques
- Provided pareto curve methodology for leakage and delay

Future Work

- Analysis of the other techniques using more general circuit, such as full adder and SRAM.
- Analysis of different set of knobs like placement location and voltage threshold.
- Evaluation of new scaling technologies (16nm) or using other Predictive Technology Model.
- Use different cost formulas like complexity